Subterranean Biology 49: 75–95 (2024)
doi: 10.3897/subtbiol.49.128521
https://subtbiol.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
Subterranean Biology
Published by The International Society for Subterranean Biology

Arthropod diversity in shallow subterranean habitats of the Appalachian Mountains

G. T. Harrison[1], Howard P. Dunleavy[1], Luisa F. Vasquez-Valverde[1], Alejandro I. Del Pozo-Valdivia[1], Kaloyan Ivanov[2], Paul E. Marek[1]

1 *Virginia Tech, Department of Entomology, Blacksburg, VA 24061, USA* **2** *Virginia Museum of Natural History, Department of Recent Invertebrates, Martinsville, VA 24112, USA*

Corresponding author: Paul E. Marek (paulemarek@gmail.com)

Academic editor: Matthew L. Niemiller | Received 6 June 2024 | Accepted 26 July 2024 | Published 14 August 2024

https://zoobank.org/82730D2A-C1BC-48BB-A0C9-800ADA028FE1

Citation: Harrison GT, Dunleavy HP, Vasquez-Valverde LF, Del Pozo-Valdivia AI, Ivanov K, Marek PE (2024) Arthropod diversity in shallow subterranean habitats of the Appalachian Mountains. Subterranean Biology 49: 75–95. https://doi.org/10.3897/subtbiol.49.128521

Abstract

Subterranean arthropods are important components of soils and contribute essential food-web functions and other ecosystem services, however, their diversity and community composition has scarcely been assessed. Subterranean pitfall traps are a commonly used method for sampling soil habitats in Europe but have never been widely implemented in the Americas. We used subterranean pitfall traps to sample previously unsurveyed arthropod communities in southwestern Virginia, U.S. Traps were placed in shallow subterranean habitats (SSHs), underground habitats close to the surface where light does not penetrate, and more specifically at the interface between the soil and underlying "milieu souterrain superficiel"—a microhabitat consisting of the air-filled interstitial spaces between rocks (abbreviated MSS). In total, 2,260 arthropod specimens were collected constituting 345 morphospecies from 8 classes, 33 orders, and 94 families. A region of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene was amplified and sequenced, and objective sequence clustering of 3% was used to establish molecular operational taxonomic units (mOTUs) to infer observed species richness. In all, 272 COI barcodes representing 256 mOTUs were documented for rare soil-dwelling arthropod taxa and are published to build a molecular library for future research in this system. This work is the first taxonomically extensive survey of North American soil-dwelling arthropods greater than 10 cm below the soil surface.

Keywords

Completeness, DNA, endogean, epigean, hypogean, milieu souterrain superficiel, MSS, shallow subterranean habitats, SSH, survey

Introduction

The need to study and describe global biodiversity has never been more urgent. Anthropogenic habitat loss has been implicated as the major driver of the currently ongoing sixth great mass extinction event in geological history (Cowie et al. 2022). Biodiversity is at greatest risk in highly diverse regions known as biodiversity hotspots (Myers et al. 2000; Hamilton et al. 2022). The Appalachian Mountains of eastern North America constitute one such hotspot, and has experienced tremendous habitat loss from exploitative mineral extraction, timber harvesting, and other land conversion practices (Stein 2000).

According to a recent review assessing global declines of arthropod diversity and abundance due to habitat loss and other factors, twice as many species show long-term population declines as those exhibiting population increases (Sánchez-Bayo and Wyckhuys 2020). Subterranean arthropods, and other low-mobility invertebrates, are at elevated risk of extinction due to high rates of endemism and physiological constraints on dispersal (Mammola et al. 2019). Arthropods constitute the most species-rich animal group on Earth, with an estimated 7 million species, of which 5.5 million belong to the class Insecta (Santos et al. 2020). The Catalog of Life (2022) documents 1,128,168 species of arthropods, making up only 16% of the species estimated to exist. As such, arthropods constitute a globally understudied portion of biodiversity at high risk of species loss and anonymous extinction, a process in which a species is lost before it is discovered and described (Lobl et al. 2024). In the face of global decline, assessing and understanding insect and arthropod diversity is of paramount importance.

Research of subterranean organisms in North America has been dominated by taxonomically narrow studies with singular focal species or groups and have yielded important discoveries that hint at a significant but yet hidden diversity (e.g. Derkarabetian et al. 2010; Harden et al. 2024a, 2024b). General sampling and taxonomic assessment of soil-dwelling arthropods in North America has rarely, if ever, been conducted greater than 10 cm below the soil surface. As a result, little to nothing is known about the diversity and taxonomic composition of the arthropod communities that occupy the shallow subterranean habitats (SSH) of North America, and especially those within the Appalachian biodiversity hotspot. Broadly defined, SSH are underground habitats close to the surface where light does not penetrate, and at a basic level include soil and underlying milieu souterrain superficiel (MSS), a microhabitat consisting of the air-filled interstitial spaces between rocks (Mammola et al. 2016). But more broadly SSH can include underground aquatic interstitial habitats (e.g. hyporheic and hypotelminorheic zones), lava tubes, epikarst, and calcrete aquifers, which are defined in Culver and Pipan (2014). These habitats (except for some lava tubes) are typically close to the surface and are variable in depth but on average extend 0.1–5 meters in depth and up to 10 m (Culver and Pipan 2014). The arthropods that live the SSH have morphological and physiological adaptations such as the lack of pigmentation and eyes, shortened legs and elongate flexible bodies (Marek et al. 2021). SSH can be considered ecotones as they serve as transitional zones between adjacent epigean (above

ground) and hypogean habitats (Prous 2004; Novak 2014). The term "hypogean" is often used to describe taxa or habitats associated with the rocky substrate that lays below the MSS, including cavernous karst cavities and cave systems distant from the surface (Prous 2004). Hypogean taxa have partially overlapping morphological adaptations to those in the SSH such as pigment and eye reduction, but contrast in having elongated appendages (Deharveng et al. 2024). Establishing tangible distinctions between different subterranean habitats is difficult as they exist more as a gradient than as distinct zones, thereby further contributing to the challenges of characterizing their diversity (Mammola et al. 2017). For the purposes of this study, SSH is used to refer to aphotic soil and MSS habitats 10–67 cm below the soil surface, but not extending into the bedrock, bedrock MSS, or into cave systems.

Sampling subterranean habitats is physically challenging, contributing to our lack of taxonomic and ecological knowledge in these systems. Although the number of studies focusing on subterranean arthropod communities has increased in recent decades, subterranean taxa continue to remain underrepresented in most biodiversity surveys due to difficulties associated with sampling the vertically distinct and ecologically important soil and MSS layers. Most studies focusing on soil invertebrate biota have taken place in the tropical and subtropical areas of the world and have focused on sampling ants (Wong and Guénard 2017) while most MSS studies have taken place in Europe, Japan, and Australia (Mammola et al. 2016; Ledesma et al. 2019; Halse and Pearson 2014). To our knowledge, SSHs have never been generally sampled in North America using ad hoc subterranean sampling.

Due in part to this lack of study, North American subterranean arthropods are vastly understudied. Essential research on them such as species descriptions, identification, and the assessment of biodiversity currently suffer from taxonomic shortcomings in which time, labor, and specialized expertise are direly required (Hebert et al. 2003; Ball and Armstrong 2011; Meierotto et al. 2019). Consequently, they are underrepresented in faunistic and ecological studies, and in major genetic repositories such as the National Center for Biotechnology Information (NCBI) and the Barcode of Life Data System (BOLD). The use of DNA barcodes, when coupled with traditional, morphology-based taxonomy, may help to partially overcome these shortcomings by streamlining identification and biodiversity assessment (Blaxter 2004). However, the taxonomic potential of DNA barcoding cannot be reached for North American soil-dwelling arthropods without the fuller representation of NCBI with expertly identified vouchers representing these taxa.

The aim of this study was to survey the subterranean arthropod communities of previously unexplored SSH in southwestern Virginia, U.S. This work also sought to generate high quality DNA barcodes for the sampled taxa in order to expand a molecular foundation for future research in North American subterranean systems. The COI barcodes generated from this study will contribute to the representation of subterranean arthropod genetic data at NCBI, Ecdysis, Global Biodiversity Information Facility (GBIF) and other biodiversity data repositories, and will serve as a tool for future work in characterizing and understanding subterranean arthropod communities of Appalachia and North America more generally.

Materials and methods

Sampling design and study sites

The subterranean pitfall traps (hereafter subterranean traps, or traps) used in this study follow the design of López and Oromí (2010), which was selected due to its efficiency in sampling subterranean arthropods, and the semi-permanent nature of the installation. The main outer shell of the trap is constructed of a polyvinyl chloride (PVC) pipe perforated with holes for entry of subterranean arthropods. The trap can remain in place, while an internal collecting receptacle itself is removable (Fig. 1). This allows for repetitive sampling without removal of the entire trap and lessens disturbance to the habitat between sampling events, thus reducing the impact of installation on the communities being sampled. The collecting receptacle is composed of a thinner plastic (polypropylene) cup with a smaller plastic (polyethylene) bait chamber secured in the center (Fig 1). Small rubber gaskets, stainless steel washers, and small bolt-nut combinations were used to stabilize the bait chamber within the center of the collecting receptacle. The bait chamber is capped with a perforated lid to allow the bait odors to disperse. Three strings are attached to the collecting receptacle, allowing it to be lowered to the bottom of the trap and then retrieved during collecting events. The perforations in the main outer shell (d = 7.95 mm) allow individuals to enter the trap, but restrict the entrance of vertebrates (e.g., small mammals, reptiles), and begin 10 cm below the soil surface to decrease the likelihood of epigean individuals entering the traps.

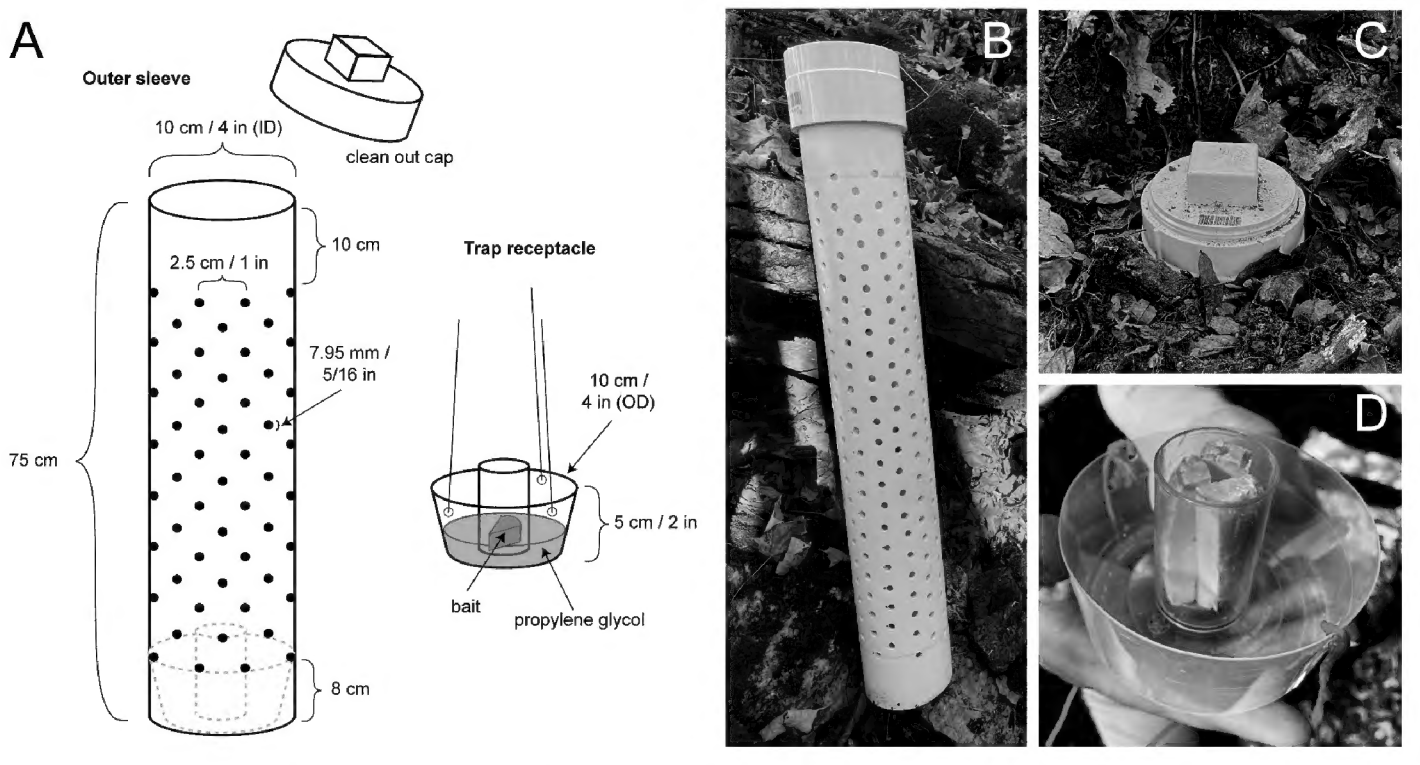

Figure 1. Design of the subterranean pitfall traps used in this study **A** subterranean pitfall trap main outer shell (left) and collecting receptacle (right) with dimensions in centimeters (cm) and inches (in). ID = inside diameter. OD = outside diameter **B** main outer shell **C** buried trap with clean out cap visible above the soil **D** collecting receptacle with bait, Limburger cheese, prior to filling with preservative. (Both metric and imperial measurements provided for some dimensions due to U.S. sourcing of materials, e.g., PVC pipe and 5/16 in. drill bit.).

The trap design allows for sampling within a range of 10–67 cm below the soil surface. The bait chamber receptacle was baited with Limburger cheese and the trap collecting receptacle filled with propylene glycol for specimen preservation. Limburger cheese was selected as bait due to its strong odor and propensity for attracting arthropods (Schneider and Culver 2004). Propylene glycol was selected due to its DNA preservation, thermal, and moisture buffering qualities (López and Oromí 2010).

A total of 20 subterranean pitfall traps were installed across three sites in southwestern Virginia (Table 1). Each site was forested, and traps were installed where evidence of frequent human disturbance was absent. Traps 1–5 were installed near the university's dolomite quarry ("Quarry"; Table 1) and traps 6–10 were installed on university land near Blacksburg ("Fallam"; Table 1). These sites are located within the Ridge and Valley ecoregion of Virginia where the rock substrate is mainly composed of sedimentary rock (Omernik 1995; U.S. Geologic Survey 2021). Dolomite was the rock type most frequently encountered while installing these traps. Traps 11–20 were installed in Floyd County, Virginia ("Starroot"; Table 1). This site is located within the Blue Ridge ecoregion of Virginia with the rock substrate heavily composed of metamorphic rock (U.S. Geologic Survey 2021). Quartzite was the rock type most frequently encountered while installing the Starroot traps. All sites were located within mixed forests dominated by oaks (*Quercus* spp.), with occasional pines (*Pinus* spp.), American beech (*Fagus grandifolia* Ehrh.), and maples (*Acer* spp.) scattered throughout (Virginia Department of Conservation and Recreation 2021). All sites were selected because they offered a topographical variety and evidence of an underlying rocky substrate consistent with the MSS that aligned with the optimal conditions described by López and Oromí (2010).

Table 1. Summary information of the study sites and traps.

Site	Trap #	Latitude, Longitude	Elevation (m)	County	GPS Accuracy (m)
Quarry	1	37.2231, -80.3832	631	Montgomery	3
	2	37.2229, -80.3799	610	Montgomery	6
	3	37.2228, -80.3818	639	Montgomery	5
	4	37.2225, -80.3875	667	Montgomery	9
	5	37.2232, -80.3836	619	Montgomery	2
Fallam	6	37.2124, -80.6093	560	Montgomery	3
	7	37.2127, -80.6085	558	Montgomery	34
	8	37.2132, -80.6054	600	Montgomery	3
	9	37.2133, -80.6049	592	Montgomery	3
	10	37.2119, -80.6090	559	Montgomery	4
Starroot	11	36.9656, -80.4185	751	Floyd	6
	12	36.9663, -80.4177	773	Floyd	3
	13	36.9685, -80.4171	771	Floyd	9
	14	36.9684, -80.4176	778	Floyd	3
	15	36.9673, -80.4170	780	Floyd	3
	16	36.9670, -80.4171	783	Floyd	5
	17	36.9668, -80.4172	781	Floyd	3
	18	36.9657, -80.4189	776	Floyd	3
	19	36.9651, -80.4186	790	Floyd	3
	20	36.9639, -80.4184	772	Floyd	3

There was slight variation in soil and rock composition between sites as well as within sites. Traps were placed in MSS microhabitats with varying sizes and quantities of rocks. Each trap was loaded with bait and preservative and operated for two weeks before the specimens were collected. A handheld Garmin eTrex 10 global positioning system (GPS) was used to record the geographical coordinates of the traps with positional accuracy recorded in meters. Traps were set on 28 December 2021 (winter) and again on 1 June 2022 (spring) for a total of four weeks of baited collection time.

Morphospecies

Specimens were removed from the propylene glycol with a sieve and pooled by order for each trap, stored in 8.0 mL Sarstedt vials with 100% ethanol, and subsequently identified to morphospecies using a Leica M125 stereomicroscope (Leica, Wetzlar, Germany). Morphospecies are operational taxonomic units identified by examination of easily observable morphological characters (Derraik et al. 2010). Because morphospecies are determined solely based on morphology, different life stages of a holometabolous insect species can be designated as separate morphospecies. One to two specimens were selected as representatives of each morphospecies. An individual of each morphospecies was photographed in ethanol in a Z-stack of 7–12 focal planes with a Canon EOS 6D SLR camera equipped with a Canon MP-E 65 mm macro lens and mounted on a Visionary Digital Passport portable imaging system (Canon, Tokyo, Japan; Visionary Digital, Charlottesville, Virginia). Helicon Focus (HeliconSoft, Kharkiv, Ukraine) was used to integrate the focal stacks into a single high resolution composite image. Each morphospecies was imaged from dorsal, ventral, and lateral views. Where appropriate, laterally compressed or coiled specimens were imaged only from lateral perspectives (e.g., some Diplopoda and Chilopoda).

Morphospecies from different sites, even those suspected to be the same species based on morphological similarity, were treated as distinct and unique in order to capture potential cryptic species. Each morphospecies was identified to at least the family level using various resources (Stehr 1991; Goulet and Huber 1993; Arnett and Thomas 2000; Arnett et al. 2002; Government of Canada 2002a, b; Triplehorn et al. 2005; Whitfield et al. 2014). Exceptions included immature mite, dipteran, coleopteran, and hemipteran specimens for which morphological identification resources do not exist, or were not accessible. Identifications of these morphospecies were retained at the order level.

DNA sequencing, barcode generation and analysis

DNA was extracted from each morphospecies using a DNeasy (Qiagen) Blood & Tissue extraction kit. The extraction protocol was modified to be less destructive, keeping specimens largely intact for morphological identification, deposition as vouchers, and potential species description (Gilbert et al. 2007). Rather than homogenizing the specimen by grinding body parts in buffer, a single puncture was made in the cuticle

with a flame-sterilized pin and the specimen transferred to the microcentrifuge tube along with the buffer solution. This puncture allows the Qiagen lysis buffer to access the softer tissues within the body without grinding the specimen. The specimen was then recovered from the DNeasy minicolumn following the final buffer wash and stored in an 8.0 mL Sarstedt vials with 100% ethanol. Specimens are deposited in the Virginia Tech Insect Collection (https://collection.ento.vt.edu) under the specimen codes provided in Suppl. material 1: table 1.

A fragment of the cytochrome *c* oxidase subunit I (COI) mitochondrial gene region was amplified utilizing polymerase chain reaction (PCR) employing the primers LCO1490 and HCO2198 (Folmer et al. 1994). These primers were selected as they have been shown to be useful for the amplification of the same COI gene region from a broad diversity of arthropods (Folmer et al. 1994; Elbrecht et al. 2019). In addition, the region corresponds to the often-used barcode region that is ubiquitous in genetic databases such as NCBI and BOLD. The PCR protocol was conducted according to Means et al. (2021a, b). Cleaning, quantification, normalization, and sequencing of amplicons on an Applied Biosystems ABI 3730 capillary sequencer was carried out by Arizona Genetics Core (University of Arizona). The raw forward and reverse chromatograms were assembled into consensus sequences in Mesquite (Version 3.61) by base calling, trimming, and quality assessment using the sequence analysis module Chromaseq (Version 1.52) with the software PHRED and PHRAP (Ewing et al. 1998; Maddison and Maddison 2009; Maddison and Maddison 2021). Sequence quality control and assembly was carried out according to the methods outlined in Vasquez-Valverde and Marek (2022) and a consensus COI sequence approximately 500–600 base-pairs in length was generated for each morphospecies. For the chromatograms that did not contig in Mesquite, additional assembly attempts were made using Geneious and various assembler algorithms (Geneious Version 2022.1.1, Build 2022-03-15). No additional contigs were retrievable. To generate the completed DNA barcodes, the consensus sequences were put in the correct reading frame within Mesquite using the sequence processing tools Reading Frame and Codon Position and aligned within Geneious using MAFFT (Version 7.490; Katoh and Standley 2013). Absence of stop codons were confirmed by visual inspection of chromatograms and examination of any erroneous single nucleotides in Mesquite, followed by designation of the proper reading frame orientation after each nucleotide check.

The barcodes were matched with existing records using GBIF's sequence-id engine querying BOLD (GBIF 2023). An additional local Basic Local Alignment Search Tool (BLAST) analysis was conducted by downloading all arthropod COI sequence data uploaded to NCBI as of 7 March 2023 and utilizing the custom batch BLAST feature in Geneious. BLAST hits with a percent identity match of ≥ 97% were accepted as molecular identifications at the species level. Hits with a percent identity match below 97%, but ≥ 95%, were accepted at the genus level (Srivathsan et al. 2022). Searches with no matches at these thresholds were left with morphological identifications at the family level. Sequences were then clustered at 3% dissimilarity (97% similarity by nucleotide identity) using objective clustering to establish molecular operational taxonomic units (mOTUs) for species binning (Meier et al. 2006; Srivathsan 2022). The divergence

threshold of 3% was employed in accordance with the original methods for objective clustering as well as similar arthropod diversity studies employing objective clustering (Hebert et al. 2003; Smith et al. 2005; Meier et al. 2006; Srivathsan et al. 2022).

Sampling completeness

Sample completeness curves with 95% confidence intervals were generated for each of the three sites in iNEXT Online (Hsieh et al. 2016) employing 1,000 bootstrap replicates to analyze the relationship between sample coverage and sample size. iNEXT Online employs Hill numbers calculated from species richness and evenness. Hill numbers estimate sample coverage through rarefaction and to extrapolate the effect of additional sampling on sample coverage (Chao et al. 2014). They represent a standardized method to equitably compare diversity across data sets even when samples originate from highly disparate natural assemblages or various techniques that they were sampled (Hsieh et al. 2016). The sample completeness curves were estimated from Hill number q = 1, or the exponential of the Shannon diversity index for each site (Shannon 1948; Hill 1973; Ellison 2010; Hsieh et al. 2016). This metric is also referred to as effective number of species (ENS). Additional sample completeness curves were generated for other Hill numbers, but are not reported here as there were no significant differences from the ENS-based curves. The two collection periods (i.e., winter and spring) were combined for the analysis.

The Starroot site was sampled with ten subterranean traps for four weeks of total collection time, while the Fallam and Quarry sites were sampled with five traps each for four weeks of total collection time, constituting equal sampling times but half the sampling effort employed at the Starroot site. To account for the unequal sampling effort across sites, extrapolated coverage estimates for double the sample size were estimated for each of the sites.

Taxonomic composition and abundance

Observed species richness and abundance were used to characterize the taxonomic composition of the samples and sites. Relative proportions of species richness by class across the three sites were analyzed as hierarchical data and visualized by treemaps in JMP Pro 16 (SAS Institute Inc., Cary, NC). Observed abundance is reported by order and includes all orders.

Results

Sampling, morphospecies, and barcode analysis

Our sampling resulted in 2,260 arthropod specimens representing 345 morphospecies (Suppl. material 1: table 1). Select morphospecies images are included in Suppl. material 1: figs 1–39, and all 717 high resolution composite images taken are deposited at Ecdysis

(https://ecdysis.org/, URLs to records in Suppl. material 1: table 2) and in the Virginia Tech Data Repository. Of the 345 morphospecies, 320 successfully amplified with COI primers, and from those amplicons, 272 successfully assembled into barcode sequences. Objective clustering revealed 256 total molecular clusters, or molecularly distinct species bins (mOTUs). Sequence dissimilarity was greater than 3% for 16 morphospecies clusters, indicating that each constituted multiple species. Each of these 16 clusters represented 2–4 species, thereby collapsing the 272 sequenced morphospecies to 256 molecularly distinct species bins. The 71 specimens that did not successfully sequence were left as morphospecies with morphological identifications only. COI barcodes were generated for the 272 sequenced specimens and deposited in NCBI (Suppl. material 1: table 1). Sequence matching via GBIF's sequence-id engine yielded 142 acceptable molecular identifications (percent identity match ≥ 95%) with 59 at the species, 40 at the genus, and 36 at the family level. One specimen first identified as an arthropod was later determined to be a pot worm (Annelida, Enchytraeidae) based on COI barcode (SPT-00078). Annelids were very abundant in the samples but only arthropods were analyzed for this study.

Sampling completeness

The sample completeness curve was highest for the Starroot site where sampling effort was double that of the other sites (Fig. 2). The asymptotic shape of the rarefied curve is indicative of the high sample coverage observed at this site (Table 2). Similarly high sample coverages were also observed for the remaining sites. The effective number of species was highest for the Quarry site and lowest for the Fallam site (Table 2). ENS was comparable between the Quarry and Starroot sites, but the observed sample completeness was higher for the Starroot site.

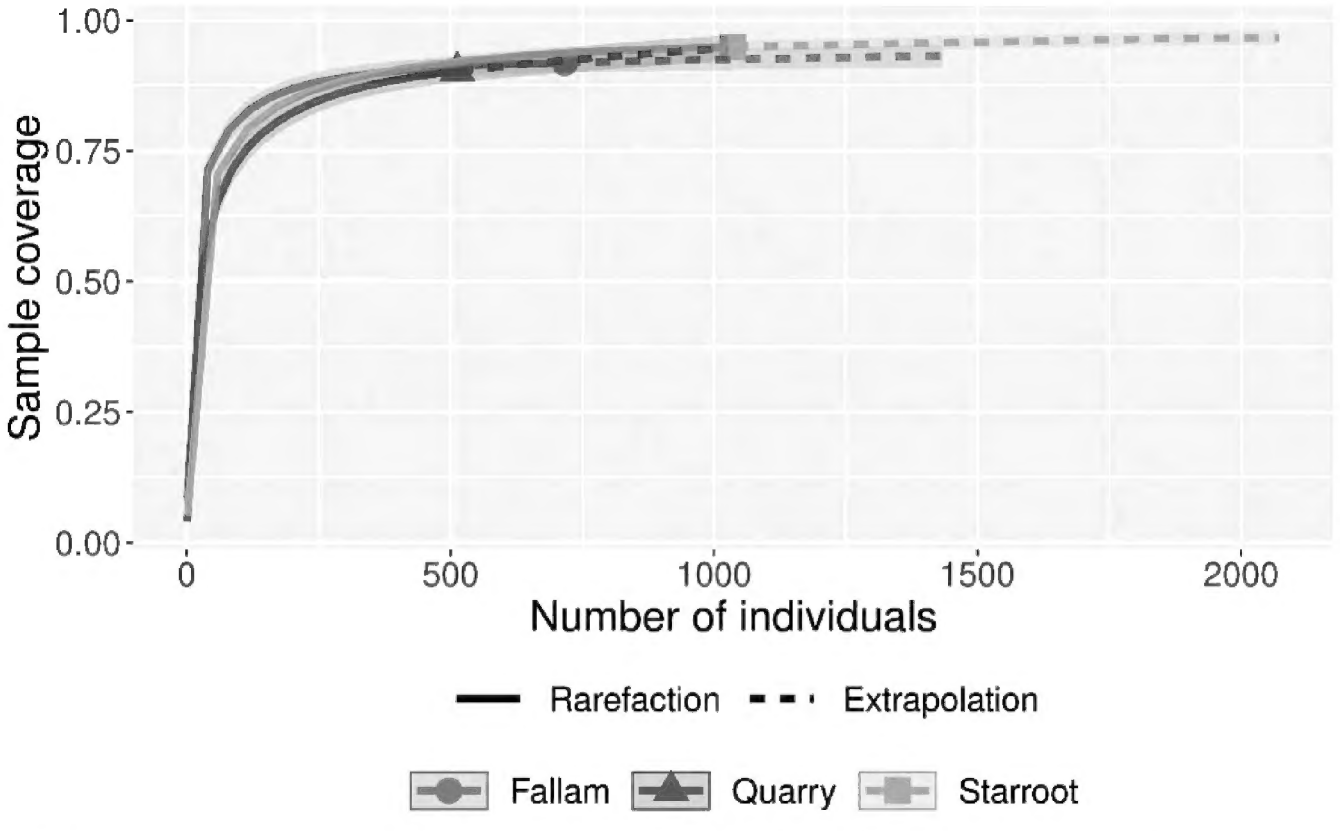

Figure 2. Sample completeness curves, with associated 95% confidence intervals, for the study sites.

Table 2. Effective number of species (ENS), Hill number q = 1, and observed and extrapolated sample coverage for each site.

Site	Quarry	Fallam	Starroot
ENS	39.485	26.882	37.691
Observed Coverage	0.903	0.919	0.949
Extrapolated Coverage	0.947	0.932	0.967

Taxonomic composition

In total, eight classes (Malacostraca, Entognatha, Insecta, Chilopoda, Diplopoda, Pauropoda, Symphyla, and Arachnida) from four arthropod subphyla (Crustacea, Hexapoda, Myriapoda, Chelicerata) were observed across all samples. Observed species richness by class was similar across the three sites (Fig. 3). The three most species rich classes across all sites were Insecta, Arachnida, and Entognatha, followed by Diplopoda, Chilopoda, and Malacostraca. Images of representative individuals from each arthropod subphylum are presented in Fig. 4 and those of each molecularly distinct species and morphospecies observed are included in Suppl. material 1: figs 1–39 [717 high resolution composite images taken are deposited at Ecdysis (https://ecdysis.org/, URLs to records in Suppl. material 1: table 2) and in the Virginia Tech Data Repository, DOI: 10.7294/26397688]. The number of morphospecies by trap, site, and collecting event are shown in Fig. 5.

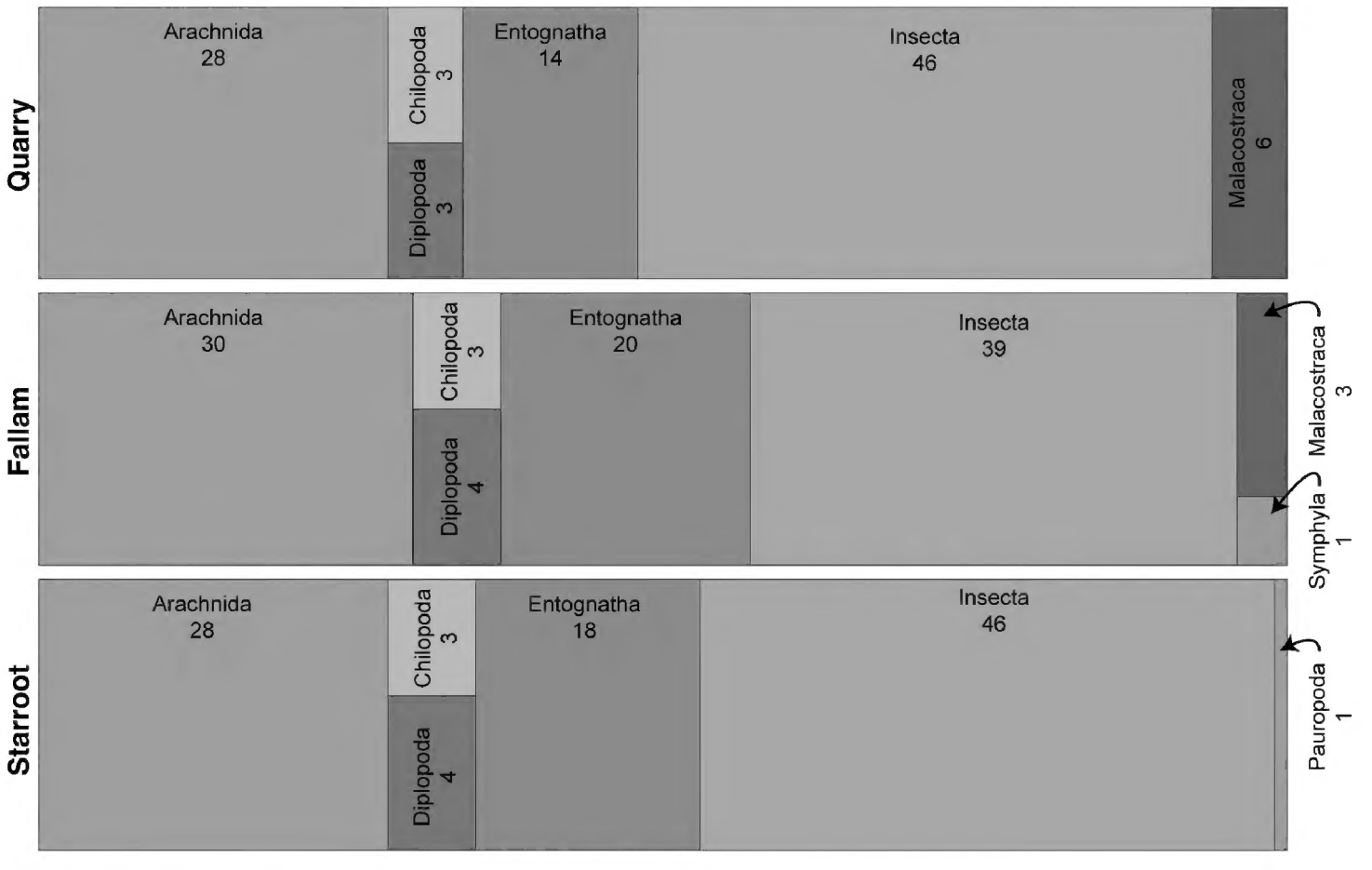

Figure 3. Observed arthropod species richness by class (reported as a percentage) across sites.

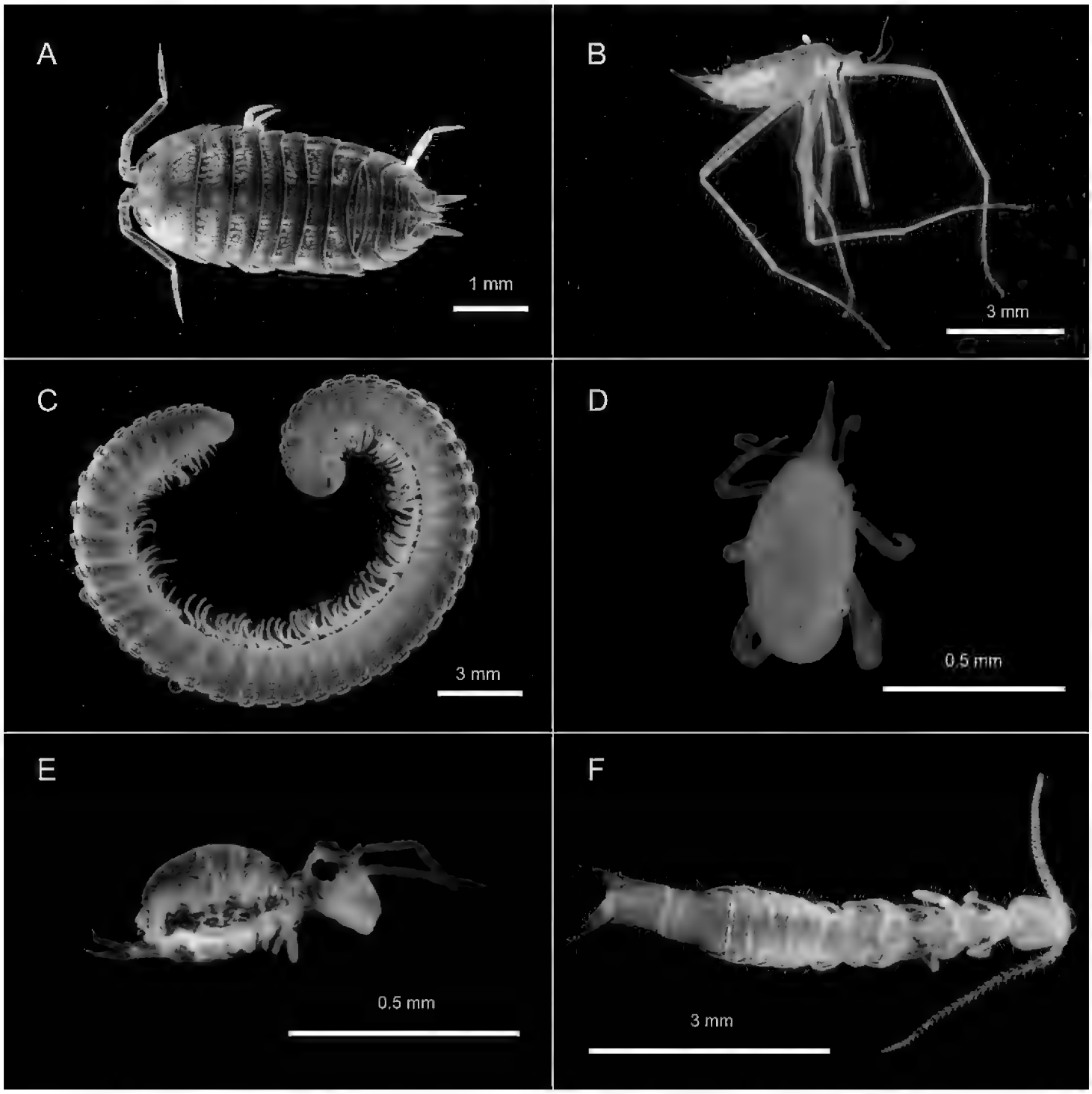

Figure 4. Representative arthropod specimens observed from each subphylum, including two from class Entognatha **A** Crustacea; dorsal view of *Cylisticus convexus*, individual SPT-0001 **B** Hexapoda; lateral view of *Chionea scita*, individual SPT-0020 **C** Myriapoda; lateral view of *Abacion* sp., individual SPT-00119 **D** Chelicerata; dorsal view of Bdellidae, individual SPT-00191 **E** Entognatha; lateral view of Sminthuridae, individual SPT-0067 **F** Entognatha; dorsal view of Japygidae, individual SPT-0061.

Abundance

Overall, the samples were dominated by hexapods and chelicerates with the orders Entomobryomorpha, Sarcoptiformes, Diptera, Hymenoptera, and Coleoptera being among the most abundant across all sites (Fig. 6). Insecta was predominantly represented by members of Diptera (40.13%), Hymenoptera (15.58%) and Coleoptera (33.12%); Arachnida by the subclass Acari or mites (66.67%); and Entognatha by the subclass Collembola (86.15%). The order Hymenoptera was predominantly represented by ants

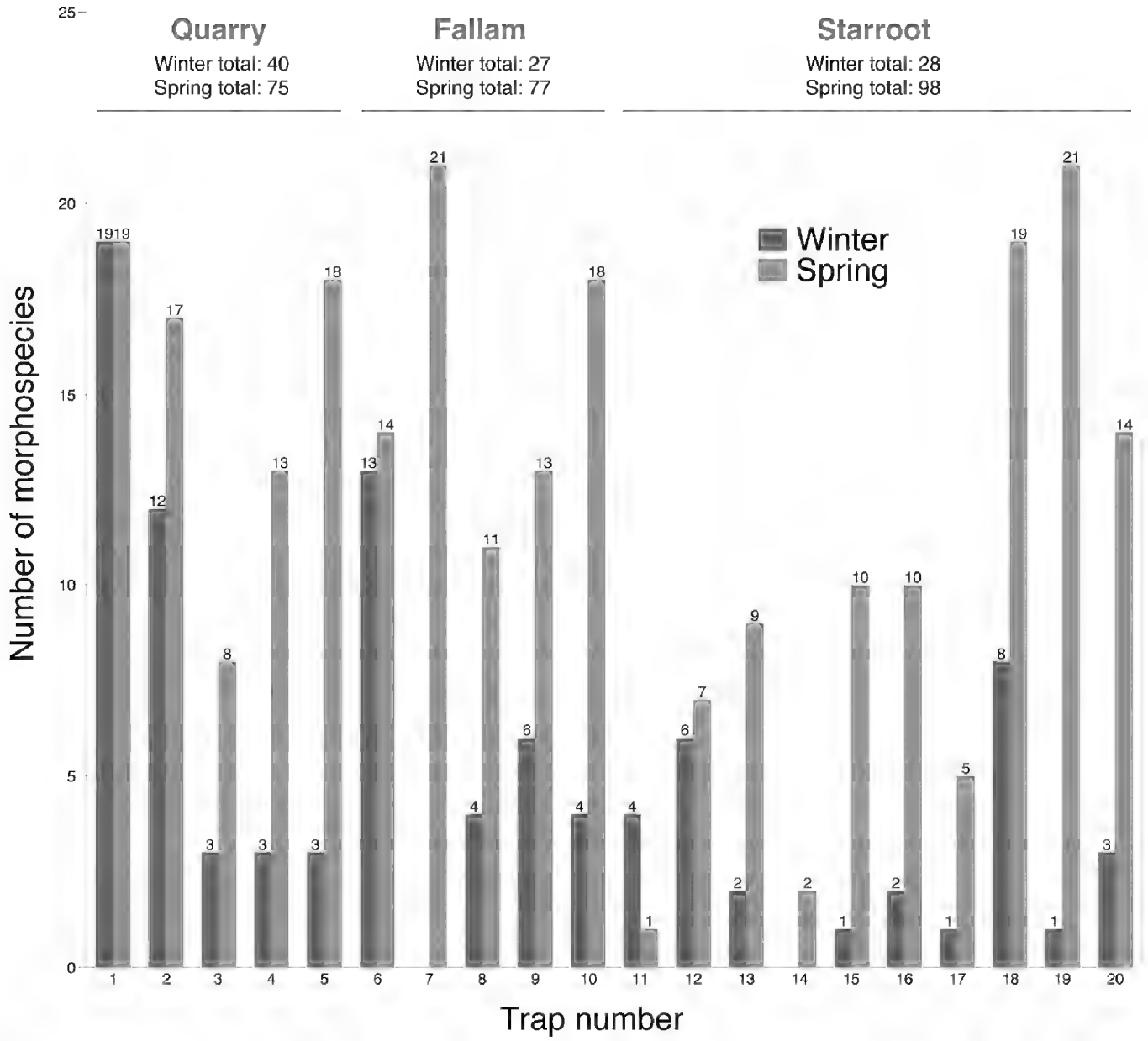

Figure 5. Number of morphospecies by trap, site, and collecting event. Data underlying this figure are in Suppl. material 1: table 3.

(Formicidae; 84.91%) across all sites. While the high abundance of springtails, mites, ants, and beetles in the collected samples is not surprising, the relatively high abundance of Diptera at the study sites can be explained by the high number of early instar larvae of the family Phoridae collected in single traps at the Quarry and Fallam sites.

Discussion

This study is the first taxonomically extensive survey of soil-dwelling arthropods in North America, and the first in the Appalachian Mountains, to employ subterranean pitfall traps capable of sampling to a depth of up to 67 cm below the soil surface. In all, 2,260 individual arthropod specimens were collected constituting 345 morphospecies and 257 molecularly distinct species (mOTUs) representing 8 classes, 33 orders, and 94 families. In total, 272 COI barcodes were sequenced and are published at NCBI. Of these, 102 constitute mOTUs that are new to the NCBI and BOLD databases. Many of the taxa recovered during the study represent new records and have not, or

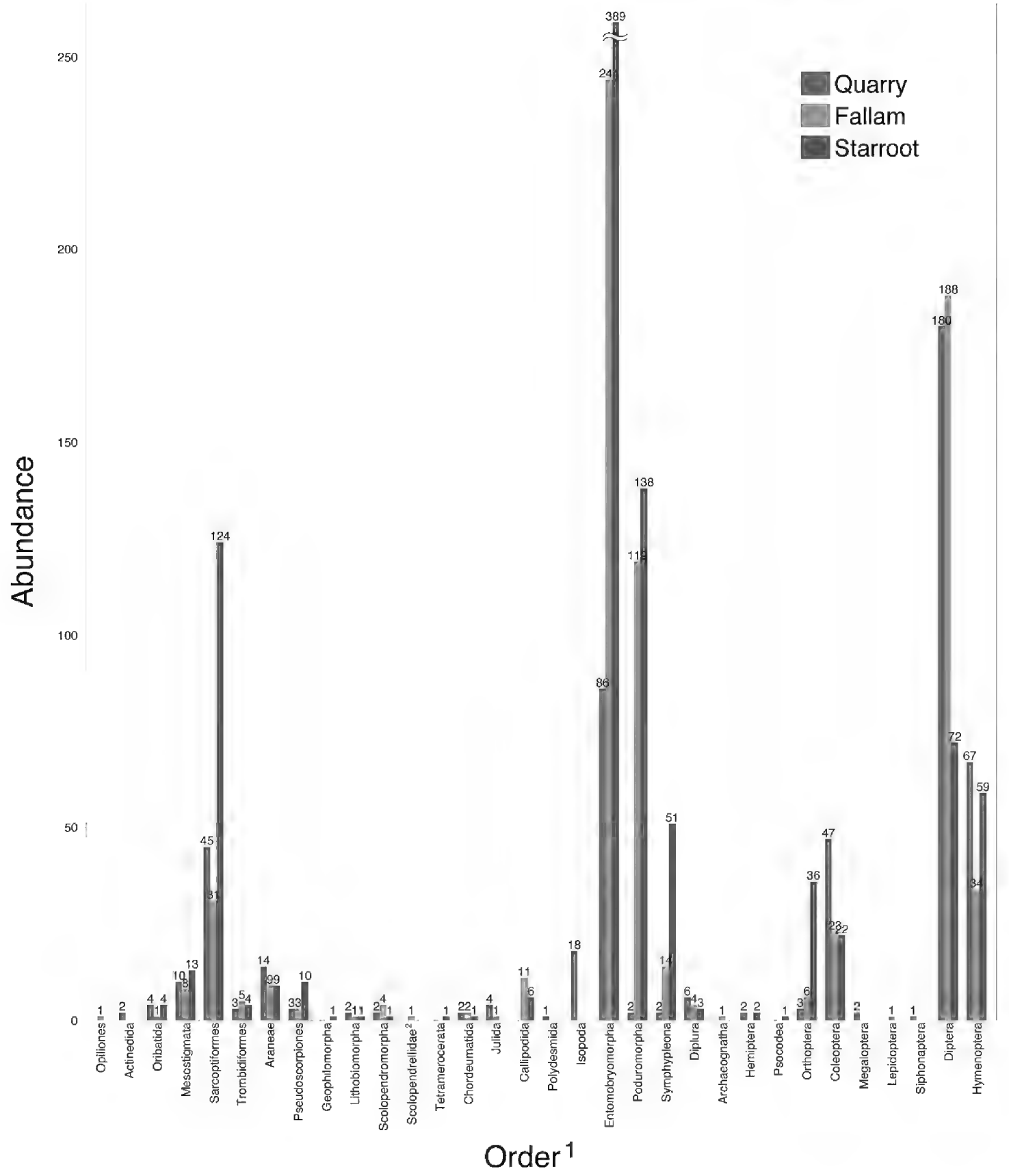

Figure 6. Arthropod orders and their abundance by site. Entomobryomorpha abundance bar for Starroot site extends off the chart and is abbreviated by double squiggly lines. ([1]Phylogenetic arrangement of orders. [2]The class Symphyla does not possess an order rank so family rank used.) Data underlying this figure are in Suppl. material 1: table 4.

have only rarely, been documented in the region. We suspect that a number of the morphospecies recovered are undescribed.

Several morphospecies exhibited hypogean/SSH adaptations: i.e., depigmentation, reduction of eyes and appendage lengthening (hypogean taxa) or shortening (SSH taxa) suggesting that they may be obligate subterranean inhabitants (Deharveng and Bedos 2018; Deharveng et al. 2024). These morphospecies included several species of ants

(Hymenoptera: Formicidae—6 out of 13 morphospecies), beetles (Coleoptera—15 of 53), flies (Diptera—18 of 62), springtails (Collembola—18 of 56), two-pronged bristletails (Diplura: Japygidae—6 of 7), spiders (Arachnida: Araneae—7 of 19), and mites (Arachnida: Mesostigmata, Sarcoptiformes, Trombidiformes—21 of 62). However, most individuals displayed morphology associated with epigean origin, which is consistent with the findings of similar studies outside North America that have shown that SSH are commonly dominated by epigean taxa (Coiffait 1958; Mammola et al. 2017). This dominance of epigean taxa might have contributed to the lower observed richness in winter as the richness and abundance of epigean arthropods, especially insects, in subterranean systems are significantly impacted by seasonality due to surface temperature influencing SSH temperature to some extent. This is not the case for hypogean taxa as climatic stability tends to increase as depth from the surface increases (Mammola et al. 2017). Alternatively, the lower observed richness in the winter may be associated with the decreased mobility of arthropods during cold and dry periods when many are in a winter diapause, or in immobile pupal stages.

The taxonomic composition of our samples is comparable to that of SSH studies in the Canary Islands (Pipan et al. 2011), Italy (Mammola et al. 2017), mainland Spain (Gilgado et al. 2014), Bulgaria (Langourov et al. 2014), and France (Juberthie 2000) with arthropods and annelids being the most abundant groups. Annelids were very abundant in our samples and were mainly represented by potworms (Enchytraeidae). Our findings differ in the detection of mollusks and crustaceans (Isopoda: Oniscidea). We only collected a single mollusk and only a few terrestrial isopods, while both groups were commonly encountered in the studies mentioned above. Our results regarding insect abundance by order align with those of Mammola et al. (2017) in that Diptera, Hymenoptera, and Coleoptera were among the most abundant groups across sites. Our findings are consistent with those of Moldenke and Lattin (1990) in that the families Phoridae, Cecidomyiidae, and Sciaridae were among the most common dipterans encountered. Diptera abundance was highly variable across sites due to, in part, the relatively large numbers of early instar Phoridae larvae observed in single traps at the Quarry and Fallam sites. Ants dominated specific traps at all sites which is consistent with the findings of Mammola et al. (2017). Where ants dominated a trap, relatively few other insects were found, and rarely were two or more ant species found in the same trap. Regarding Coleoptera richness by family, our findings are consistent with those of Moldovan (2005) who reported ground beetles (Carabidae; 14.3% of total abundance in our samples) to be less common than round fungus beetles (Leiodidae; 18.4%). This is in contrast with the study of Mammola et al. (2017), in which carabids were largely observed as larvae with only a single mature specimen collected. In contrast to the studies mentioned above, rove beetles (Staphylinidae; 24.5%) were the most species rich beetle group present in our samples. Hemipterans were rare, consistent with the findings of Langourov et al. (2014). This may be due to bait-bias as phytophagous hemipterans are likely less attracted to protein-based bait than their coleopteran equivalents. Bait-bias has been shown to influence the abundance and richness of taxa recovered in previous surveys of arthropods (Checa et al. 2019). Orthoptera was among the most abundant insect orders at the Starroot site, largely due to

the relatively high abundance of the locally common *Ceuthophilus guttulosus*, F. Walker, 1869 (Orthoptera, Raphidophoridae).

Little overlap in molecularly distinct species was observed between the different sites with the Quarry and Fallam sites sharing two species (Hymenoptera: *Stenamma schmittii* Wheeler, 1903; Diptera: *Triphleba aequalis* Schmitz, 1919), and the Quarry and Starroot sites sharing two species (Hymenoptera: *Stenamma schmittii* Wheeler, 1903; Araneae: *Cicurina pallida* Keyserling, 1887). The ant species *Stenamma schmittii* Wheeler, 1903 was observed across all sites. This lack of pronounced species overlap is consistent with Lamoncha (1994) who recorded minimal overlap of oribatid mites between nearby subterranean sites in North and South Carolina, U.S. This trend is consistent with the notion that subterranean arthropod communities are often highly specific to the individual habitats they occupy (Menta and Remelli 2020). That said, these patterns may be a byproduct of insufficient sampling and further work is needed to support these trends, especially for sites with lower sampling effort.

The high sample coverage at the Starroot site suggests that future surveys of similar experimental design should employ ≥ 10 subterranean traps per site and ≥ 4 weeks total of baited-collection time. Further studies are needed to optimize the application of subterranean pitfall traps in North America. Subsequent investigations may also address other factors including presence or choice of baits, duration of sampling, and digging-in effects. Previous work has shown taxon bias associated with baits (Checa et al. 2019) and recency of placement (Mammola et al. 2017). Prior work has shown that baited subterranean traps preponderantly select for epigean taxa such as ants and phorid flies (Mammola et al. 2017). Similarly, epigean taxa may be more abundant in freshly excavated traps and these digging-in effects have been shown in aboveground pitfall traps but not demonstrated in subterranean traps. Future studies without baits and/or including a post excavation and installation settling period may reduce epigean bias and narrow sampling to hypogean (broadly defined) taxa.

Conclusions

Our findings suggest that the understudied SSH arthropod communities of the Appalachian Mountains, and North America, are highly diverse and warrant further study. This is consistent with the findings of several studies showing high arthropod diversity within subterranean assemblages both more broadly and specifically within particular groups such as beetles, spiders, and mites (Lamoncha 1994; Christman et al. 2005; Mammola et al. 2017; Ledesma et al. 2019). The close, but not quite-asymptotic nature of the sample completeness curves for each of our sites indicate that more species remain undocumented, and additional sampling is needed to better characterize subterranean arthropod diversity of southwestern Virginia and the Appalachian region. We still lack solid taxonomic, ecological, and molecular foundations for our understanding of the arthropod communities that inhabit the North American SSH. It is our hope that the primary taxonomic and barcode data generated through this work will serve as a foundation upon which future studies of North American subterranean systems may build upon.

Acknowledgements

This work was supported by a National Science Foundation grant to P. Marek (Division of Environmental Biology # 1916368). We thank Ruth Neumann, also known as the artist Starroot, of Floyd County, Virginia, for permission to sample arthropods at her home (https://starroot.com/). We thank Barry Keith for assistance and expertise with trap fabrication, and India Garden Restaurant (Blacksburg, VA) for providing deli cups for fabrication of extra collecting receptacles. Kayla Elahi provided helpful assistance with uploading specimen metadata to Ecdysis. We thank Lindi Fitzpatrick and Maggie Thomas for assistance with specimen collection and processing. We are grateful to Amrita Srivathsan for assistance with Python code for objective clustering, and Steph Benbow for assistance with field work and figure preparation for this manuscript. We appreciate the comments of reviewers Stefano Mammola and David Culver who improved previous versions of the manuscript.

References

Arnett RH, Thomas MC (2000) American Beetles. Volume I, Archostemata, Myxophaga, Adephaga, Polyphaga. CRC Press, Boca Raton, Fl, 464 pp. https://doi.org/10.1201/9781482274325

Arnett RH, Thomas MC, Skelley PE, Howard Frank J (2002) American Beetles. Volume II, Polyphaga: Scarabaeoidea through Curculionoidea. CRC Press, Boca Raton, Fl, 880 pp. https://doi.org/10.1201/9781420041231

Balestra,V, Lana E, Carbone C, De Waele J, Manenti R, Galli L (2021) Don't forget the vertical dimension: assessment of distributional dynamics of cave-dwelling invertebrates in both ground and parietal microhabitats. Subterranean Biology 40: 43–63. https://doi.org/10.3897/subtbiol.40.71805

Ball S, Armstrong K (2011) DNA barcodes for insect pest identification: A test case with tussock moths (Lepidoptera: Lymantriidae). Canadian Journal of Forest Research 36: 337–350. https://doi.org/10.1139/x05-276

Blaxter ML (2004) The promise of a DNA taxonomy. Philosophical Transactions of the Royal Society of London Biological Sciences 359(1444): 359669–359679. https://doi.org/10.1098/rstb.2003.1447

Cardoso P, Barton PS, Birkhofer K, Chichorro F, Deacon C, Fartmann T, Fukushima CS, Gaigher R, Habel JC, Hallmann CA, Hill MJ, Hochkirch A, Kwak ML, Mammola S, Noriega JA, Orfinger AB, Pedraza F, Pryke JS, Roque FO, Settele J, Simaika JP, Stork NE, Suhling F, Vorster C, Samways MJ (2020) Scientists' warning to humanity on insect extinctions. Biological conservation 242: 108426. https://doi.org/10.1016/j.biocon.2020.108426

Catalog of Life [COL] (2022) Catalog of Life subgroup Arthropoda COL | Arthropoda. https://www.catalogueoflife.org/data/taxon/RT [April 13, 2023]

Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs 84: 45–67. https://doi.org/10.1890/13-0133.1

Checa MF, Donoso DA, Rodriguez J, Levy E, Warren A, Willmott K (2019) Combining sampling techniques aids monitoring of tropical butterflies. Insect Conservation and Diversity 12: 362–372. https://doi.org/10.1111/icad.12328

Christman MC, Culver DC, Madden MK, White D (2005) Patterns of endemism of the eastern North American cave fauna. Journal of Biogeography 32: 1441–1452. https://doi.org/10.1111/j.1365-2699.2005.01263.x

Coiffait H (1958) Les coleopteres du sol. Vie et Milieu Supplement 7: 1–204.

Cowie H, Bouchet P, Fontaine B (2022) The sixth mass extinction: fact, fiction or speculation? Biological Reviews 97(2): 640–663. https://doi.org/10.1111/brv.12816

Culver DC, Pipan T (2014) Shallow subterranean habitats: ecology, evolution, and conservation. Oxford University Press, New York & London, [xxii +] 258 pp. https://doi.org/10.1093/acprof:oso/9780199646173.001.0001

Deharveng L, Bedos A (2018) Diversity of terrestrial invertebrates in subterranean habitats. In: Moldovan OT, Kováč L, Halse S (Eds) Cave Ecology. Springer, Ecological Studies 235, 541. https://doi.org/10.1007/978-3-319-98852-8_16

Deharveng L, Bedos A, Pipan T, Culver DC (2024) Global subterranean biodiversity: A unique pattern. Diversity 16(3): 157. https://doi.org/10.3390/d16030157

Derkarabetian S, Steinmann DB, Hedin M (2010) Repeated and time-correlated morphological convergence in cave-dwelling harvestmen (Opiliones, Laniatores) from montane western North America. PLoS ONE 5(5): e10388. https://doi.org/10.1371/journal.pone.0010388

Derraik JGB, Early JW, Closs GP, Dickinson KJM (2010) Morphospecies and taxonomic species comparison for hymenoptera. Journal of Insect Science 10: 108. https://doi.org/10.1673/031.010.10801

Elbrecht V, Braukmann TWA, Ivanova NV, Prosser SWJ, Hajibabaei M, Wright M, Zakharov EV, Hebert PDN, Steinke D (2019) Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ 7: e7745. https://doi.org/10.7717/peerj.7745

Ellison AM (2010) Partitioning diversity. Ecology 91: 1962–1963. https://doi.org/10.1890/09-1692.1

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Research 8: 175–185. https://doi.org/10.1101/gr.8.3.175

Folmer O, Black M, Hoeh W, Lutz R, RVijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Geneious Prime (2022) [.1.1, Build 2022.3.15] Dotmatics. https://www.geneious.com

Gilbert MTP, Moore W, Melchior L, Worobey M (2007) DNA extraction from dry museum beetles without conferring external morphological damage. PloS one 2(3): e272. https://doi.org/10.1371/journal.pone.0000272

Gilgado JD, Ledesma E, Cuesta E, Arrechea E, Zapata de la Vega JL, Sánchez-Ruiz A, Ortuño VM (2014) Dima assoi Pérez Arcas 1872 (Coleoptera: Elateridae): from montane to hypogean life. An example of exaptations to the subterranean environment. Annales de la Société entomologique de France 50(3/4): 264–271. https://doi.org/10.1080/00379271.2014.981421

Global Biodiversity Information Facility (2023) Sequence-id engine. [accessed, 2023] https://www.gbif.org/tools/sequence-id

Goulet H, Huber J (1993) Ottawa: Agriculture Canada Hymenoptera of the World: An Identification Guide to Families.

Government of Canada PS and PC (2002a) Manual of Nearctic Diptera Vol. 1 / coordinated by J.F. McAlpine ... [et al.].: A54-3/27E-PDF - Government of Canada Publications - Canada.ca. https://publications.gc.ca/site/eng/9.817747/publication.html [April 13, 2023]

Government of Canada PS and PC (2002b) Manual of Nearctic Diptera Vol. 2 / coordinated by J.F. McAlpine: A54-3/28E-PDF - Government of Canada Publications - Canada.ca. https://publications.gc.ca/site/eng/9 .817749/publication.html [April 13, 2023]

Halse S, Pearson GB (2014) Troglofauna in the vadose zone: comparison of scraping and trapping results and sampling adequacy. Subterranean Biology 13: 17–34. https://doi.org/10.3897/subtbiol.13.6991

Hamilton H, Smyth RL, Young BE, Howard TG, Tracey C, Breyer S, Cameron DR, Chazal A, Conley AK, Frye C, Schloss C (2022) Increasing taxonomic diversity and spatial resolution clarifies opportunities for protecting US imperiled species. Ecological Society of America: Ecological Applications 32(3): e2534. https://doi.org/10.1002/eap.2534

Harden CW, Davidson RL, Malabad TE, Caterino MS, Maddison DR (2024a) Phylogenetic systematics of the enigmatic genus *Horologion* Valentine, 1932 (Coleoptera, Carabidae, Trechinae, Horologionini), with description of a new species from Bath County, Virginia. Subterranean Biology 48: 1–49. https://doi.org/10.3897/subtbiol.48.114404

Harden CH, Caterino MS (2024b) Revision of the *elongatus* and *pecki* species groups of *Anillinus* Casey, 1918 (Coleoptera: Carabidae: Trechinae: Anillini), with four new species and notes on *Anillinus turneri* Jeannel, 1963. Insecta Mundi 1057: 1–31. https://journals.flvc.org/mundi/article/view/136788/141473

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings. Biological Sciences 270: 313–321. https://doi.org/10.1098/rspb.2002.2218

Hill, MO (1973) Diversity and evenness: a unifying notation and its consequences. Ecology 54: 427–432. https://doi.org/10.2307/1934352

Hsieh TC, Ma KH, Chao A (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution 7: 1451–1456. https://doi.org/10.1111/2041-210X.12613

Juberthie C (2000) The diversity of the karstic and pseudokarstic hypogean habitats in the world. In: Wilkens H, Culver DC, Humphreys EF (Eds) Subterranean Ecosystems. Elsevier, Amsterdam, 17–39.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30: 772–780. https://doi.org/10.1093/molbev/mst010

Lamoncha KL (1994) Spatial and Temporal Variation in Abundance and Diversity of Soil Oribatida in Southeastern Appalachian Forests. Master's Thesis, University of Georgia, Athens, GA, USA.

Langourov M, Lazarov S, Stoev P, Guéorguiev B, Deltshev C, Petrov B, Andreev S, Simov N, Bekchiev R, Antonova V, Ljubomirov T, Dedov I, Georgiev D (2014) New and interesting records of the MSS and cave fauna of Vitosha Mt., Bulgaria.

Ledesma E, Jiménez-Valverde A, Castro A, Aguado-Aranda P, Ortuño V (2019) The study of hidden habitats sheds light on poorly known taxa: spiders of the Mesovoid Shallow Substratum. ZooKeys 841: 39–59. https://doi.org/10.3897/zookeys.841.33271

Löbl I, Klausnitzer B, Hartmann M, Krell FT (2023) The silent extinction of species and taxonomists—An appeal to science policymakers and legislators. Diversity 15(10): 1053. https://doi.org/10.3390/d15101053

López H, Oromí P (2010) A type of trap for sampling the mesovoid shallow substratum (MSS) fauna. Speleobiology Notes 2: 7–11.

Maddison DR, Maddison WP (2021) Chromaseq: a Mesquite package for analyzing sequence chromatograms. Version 1.53. http://chromaseq.mesquiteproject.org

Maddison W, Maddison D (2009) MESQUITE: a modular system for evolutionary analysis. 11.

Mammola S, Cardoso P, Culver DC, Deharveng L, Ferreira RL, Fišer C, Galassi DMP, Griebler C, Halse S, Humphreys WF, Isaia M, Malard F, Martinez A, Moldovan OT, Niemiller ML, Pavlek M, Reboleira ASPS, Souza-Silva M, Teeling EC, Wynne JJ, Zagmajster M (2019) Scientists' warning on the conservation of subterranean ecosystems. BioScience 69(8): 641–650. https://doi.org/10.1093/biosci/biz064

Mammola S, Giachino PM, Piano E, Jones A, Barberis M, Badino G, Isaia M (2016) Ecology and sampling techniques of an understudied subterranean habitat: the Milieu Souterrain Superficiel (MSS). The Science of Nature 103: 88. https://doi.org/10.1007/ s00114-016-1413-9

Mammola S, Piano E, Giachino PM, Isaia M (2017) An ecological survey of the invertebrate community at the epigean/hypogean interface. Subterranean Biology 24: 27–52. https://doi.org/10.3897/subtbiol.24.21585

Marek PE, Buzatto BA, Shear WA, Means JC, Black DG, Harvey MS, Rodriguez J (2021) The first true millipede—1306 legs long. Scientific Reports 11: 1–8. https://doi.org/10.1038/s41598-021-02447-0

Means JC, Hennen DA, Marek PE (2021a) A revision of the minor species group in the millipede genus *Nannaria* Chamberlin, 1918 (Diplopoda, Polydesmida, Xystodesmidae). ZooKeys 1030: 1–180. https://doi.org/10.3897/zookeys.1030.62544

Means JC, Hennen DA, Tanabe T, Marek PE (2021b) Phylogenetic Systematics of the Millipede Family Xystodesmidae. Insect Systematics and Diversity 5(2): e1. https://doi.org/10.1093/isd/ixab003

Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA Barcoding and Taxonomy in Diptera: A Tale of High Intraspecific Variability and Low Identification Success. Systematic Biology 55: 715–728. https://doi.org/10.1080/10635150600969864

Meierotto S, Sharkey MJ, Janzen DH, Hallwachs W, Hebert PDN, Chapman EG, Smith MA (2019) A revolutionary protocol to describe understudied hyperdiverse taxa and overcome the taxonomic impediment. Deutsche Entomologische Zeitschrift 66: 119–145. https://doi.org/10.3897/dez.66.34683

Menta C, Remelli S (2020) Soil Health and Arthropods: From Complex System to Worthwhile Investigation. Insects 11: 54. https://doi.org/10.3390/insects11010054

Moldenke AR, Lattin JD (1990) Dispersal characteristics of old-growth soil arthropods: The potential for loss of diversity and biological function. Northwest Environmental Journal 6: 408–409.

Moldovan O (2005) Beetles. In: Culver D, White WB (Eds) Encyclopedia of Caves. Elsevier Academic Press, NY, 45–51.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403: 853–858. https://doi.org/10.1038/35002501

Novak T, Culver DC, Pipan T (2014) Shallow subterranean habitats. Ecology, evolution and conservation. Acta Carsologica 43: 1–2. https://doi.org/10.3986/ac.v43i2-3.1659

Omernik JM (1995) Ecoregions: A framework for managing ecosystems. The George Wright Forum 12(1): 35–50.

Pipan T, López H, Oromí P, Polak S, Culver DC (2011) Temperature variation and the presence of troglobionts in terrestrial shallow subterranean habitats. Journal of Natural History 45: 253–273. https://doi.org/10.1080/00222933.2010.523797

Prous X, Ferreira RL, Martins RP (2004) Ecotone delimitation: Epigean–hypogean transition in cave ecosystems. Austral Ecology 29: 374–382. https://doi.org/10.1111/j.1442-9993.2004.01373.x

Sánchez-Bayo F, Wyckhuys K (2020) Further evidence for a global decline of the entomofauna. Austral Entomology 60: 9–26. https://doi.org/10.1111/aen.12509

Santos J, Fernandes G, Almeida W (2020) Arthropods: Why It Is So Crucial to Know Their Biodiversity? In: Santos JC, Fernandes GW (Eds) Measuring Arthropod Biodiversity. Springer, Cham, 3–11. https://doi.org/10.1007/978-3-030-53226-0_1

Schneider K, Culver, DC (2004) Estimating subterranean species richness using intensive sampling and rarefaction curves in a high density cave region in West Virginia. Journal of Cave and Karst Studies 66(2): 39–45. https://www.caves.org/wp-content/uploads/Publications/JCKS/v66/v66n2-Schneider.pdf

Shannon CE (1948) A Mathematical Theory of Communication. Bell System Technical Journal 27: 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Smith MA, Fisher BL, Hebert PDN (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philosophical Transactions of the Royal Society B: Biological Sciences 360: 1825–1834. https://doi.org/10.1098/rstb.2005.1714

Srivathsan A (2022) Objective clustering Python code. https://github.com/asrivathsan/obj_cluster

Srivathsan A, Loh RK, Ong EJ, Lee L, Ang Y, Kutty SN, Meier R (2022) Network analysis with either Illumina or MinION reveals that detecting vertebrate species requires metabarcoding of iDNA from a diverse fly community. Molecular Ecology 32: 6418–6435. https://doi.org/10.1111/mec.16767

Stehr FW (1991) Immature insects. Vol. 2. Updated printing. Kendall/Hunt, Dubuque, Iowa, 975 pp.

Stein B, Kutner LS, Adams JS (2000) Environmental Conservation - Precious Heritage: The Status of Biodiversity in the United States. https://doi.org/10.1093/oso/9780195125191.001.0001

Triplehorn CA, Johnson NF (2005) Borror and Delong's Introduction to the Study of Insects. Brooks. 7th edn. Cengage Learning, 888 pp.

U.S. Geological Survey (2021) Digital geologic maps of the US states with consistent lithology, age, GIS database structure, and format. [Accessed October 14, 2021] https://mrdata.usgs.gov/geology/state/map-us.html

Vasquez-Valverde LF, Marek PE (2022) Phylogenetic review of the millipede genus *Cherokia* Chamberlin, 1949 (Polydesmida, Xystodesmidae). ZooKeys 1106: 141–163. https://doi.org/10.3897/zookeys.1106.81386

Virginia Department of Conservation and Recreation (2021) Overview of the physiography and vegetation of Virginia, 59 pp. https://www.dcr.virginia.gov/natural-heritage/natural-communities/document/ncoverviewphys-veg.pdf

Whitfield JB, Doyen JT, Purcell AH, Daly HV (2014) Daly and Doyen's introduction to insect biology and diversity. 3rd edn. Oxford University Press, New York, 718 pp.

Wong MKL, Guénard B (2017) Subterranean ants: summary and perspectives on field sampling methods, with notes on diversity and ecology (Hymenoptera: Formicidae). Myrmecological News 15: 1–16. https://doi.org/10.25849/myrmecol.news_025:001

Supplementary material 1

A combined file

Authors: G. T. Harrison, Howard P. Dunleavy, Luisa F. Vasquez-Valverde, Alejandro I. Del Pozo-Valdivia, Kaloyan Ivanov, Paul E. Marek

Data type: pdf

Explanation note: Morphospecies occurrences, NCBI accession numbers; morphospecies figures 1–39; morphospecies records, including occurrence information and image URLs at Ecdysis.org; data underlying Fig. 5; data underlying Fig. 6.

Link: https://doi.org/10.3897/subtbiol.49.128521.suppl1

Supplementary material 2

Darwin core occurrence file

Authors: G. T. Harrison, Howard P. Dunleavy, Luisa F. Vasquez-Valverde, Alejandro I. Del Pozo-Valdivia, Kaloyan Ivanov, Paul E. Marek

Data type: csv

Explanation note: CSV file of occurrence data in Darwin Core data format.

Link: https://doi.org/10.3897/subtbiol.49.128521.suppl2